

SUPPLEMENTAL MATERIAL

A large nosocomial outbreak of *bla*_{KPC} carbapenemase-producing *Escherichia coli* demonstrates refractory carbapenemase gene outbreaks involving sink sites requiring novel approaches to infection control

Decraene/Phan/George/Wyllie et al.

Supplementary Methods

Patient/microbiology data linkage and epidemiological analysis

Missing data

In this study, record linkage was performed using NHS numbers using Microsoft SQL Server and Python scripts. NHS numbers were present in all patient administration system records in the study period. A small number of microbiology records lacking NHS numbers could not be linked and were excluded from analysis. We investigated whether this policy was likely to result in substantial under-ascertainment of CRE isolations, by examining laboratory records of CRE isolations from patients without NHS numbers ('NHS number-less patients').

In particular, during the study period:

Using local hospital numbers to identify patients, we found there were 96 isolations of any CRE organism from 75 NHS number-less patients, corresponding to a sensitivity of detection of CRE isolation in microbiological records using NHS number alone of 98.4%. The numbers of samples isolated from NHS number-less individuals in the calendar years 2010 to 2016 were respectively 3, 8, 25, 26, 12, 11, and 11 samples.

CRE *E. coli* was isolated from 12 NHS number-less individuals. The number of samples yielding CRE *E. coli* isolated from NHS number-less individuals in the calendar years 2010 to 2016 in each year was 0, 1, 0, 2, 1, 3 and 6 respectively; the numbers isolated from patients on the Manchester Heart Centre (MHC; based on the stated location on the microbiology request) were 0, 0, 0, 0, 0, 1, 0 respectively.

In our view, this level of missingness is not likely to compromise the study analysis or conclusions.

Segmented regression analysis - additional detail

Counts of carbapenem-resistant *Enterobacteriaceae* (CRE) screens and of first positive isolations (of carbapenem-resistant *E. coli*, carbapenem-resistant *K. pneumoniae* or any CRE) two or more calendar days after admission (where the day of hospital admission is day 0) were obtained from routine diagnostic laboratory data for the period from week 3 of 2010 to week 52 of 2016.

Four periods were defined:

- Period 1 (P1): from week 3 2010 (earliest available data) to week 26 2014 (after which a new screening policy was implemented in response to the publication of national CPE guidance)
- Period 2 (P2): from week 27 2014 to week 39 2015 (after which an intervention which included closure and replumbing of wards W3/W4 began; reference period for the analysis)
- Period 3 (P3): from week 40 2015 to week 2 2016 (during which wards W3/W4 were closed)

- Period 4 (P4): from week 3 2016 to week 52 2016 (from re-opening of wards W3/W4 to the latest available data)

Only the first detection of CRE from a screening swab for an individual was counted, and that only when the screening swab was taken on day 2 or later of admission. Some patients were on multiple wards on the day of first detection; for such individuals, they were counted as being on the MHC if any ward was within the MHC.

In total, 184,539 screens were taken during the four periods. In patients screened on days 0 and 1 following admission, 256 patients had a first positive CRE detection from 65,975 screening swabs taken from the 61,079 patients screened (i.e. 0.4% of patients screened were positive).

In patients in hospital for two or more days, 2,445 patients had a first positive CRE detection from 118,564 screening swabs taken from the 98,751 patients screened (i.e. 2.4% of patients screened were positive).

Thus, first positivity with CRE organisms was predominantly detected in patients two or more days into admission, and we analysed this group further.

Preliminary data analysis indicated over-dispersion and a possible excess of zero counts, as well as autocorrelation. We therefore initially fitted zero-inflated negative binomial models of weekly counts of positive detections from screens, with the natural logarithm of the number of patients screened post-2-day per week as an offset. In particular, we counted a patient as screened in the week whether they had one or three screens to avoid differential counting of screened patients if more screens per week were done over time. We used P2 as the reference period because fewer screens were done in P1 (following the policy in place at the time) (Fig.S2), and therefore incidence rates could be expected to be lower in P1 than P2 and subsequently simply because fewer patients were screened per week. We used the “pscl” package (version 1.5.1) in R (version 3.4.1) to fit zero-inflated negative binomial models as mixture models, with intercept-only logistic regression models for zero counts and negative binomial models for non-zero counts. The covariates in the count model were period, as above, and location: MHC versus other ward locations (reference level). All models included interaction terms between period and location to enable comparisons between changes in incidence in MHC vs non-MHC locations.

Comparison of zero-inflated models with standard negative binomial models with the same covariates using the AIC-corrected Vuong test indicated superiority of standard negative binomial models. To account for autocorrelation of the dependent variable, standard errors of parameter estimates were adjusted using Newey and West heteroscedasticity and autocorrelation consistent covariance matrix estimators¹.

Ethical approvals

Public Health England (PHE) has permission to process confidential patient information for communicable disease outbreak surveillance/management (Health Service [Control of Patient Information] Regulations, 2002). Identifiable data was used only for dataset linkage; all analyses were undertaken using anonymised data.

Details on isolate selection for sequencing, DNA extractions, sequencing, sequence data processing and sequence data analysis methods

We used the TRACE database to identify the earliest available sequence per patient - denoted as “first CR-*E. coli*” if the stored isolate collection date was ≤ 7 days from the first isolation date in the TRACE database, or “sequential CR-*E. coli*” if the stored isolate date was after this (Table S2). KPC-EC isolates from a Public Health England (PHE) project sequencing the first ten KPC-*Enterobacteriaceae* from hospitals in northwest England (2009-2014) were also included (Table S3, denoted “regional study”). We also sequenced five KPC-EC cultured during an initial environmental prevalence survey on W3/W4 (10/Mar/2015); any KPC-EC isolated as part of subsequent, intermittent IPC-associated environmental sampling (09/Apr/2015-17/Nov/15); and isolates available at the time of analysis from environmental (n=2) and patient (n=5) samples from a separate, on-going study (commenced January 2016; Table S2, denoted “environmental”).

DNA extraction for long-read sequencing (Nanopore/PacBio)

Isolates were cultured from frozen stocks (-80°C) on Columbia Blood agar in the presence of three ertapenem discs (10µg; Oxoid, Thermo Fisher Scientific, USA) overnight at 37°C. Genomic DNA was isolated using the Qiagen Genomic-tip 100/G kit (Qiagen, Germany) following the manufacturer’s recommendations. DNA was quantified using the Qubit 2.0 Fluorometer (Life Technologies, USA) and fragment length assessed using the TapeStation 2200 (Agilent, UK).

MinION sequencing

DNA was prepared for sequencing in April 2016 using a modified version of the manufacturer’s protocol (SQK-MAP006; ONT, UK). Fragmentation was not performed and library preparation began with 4-6 µg of DNA. All bead-based clean-up steps were extended to 10 minutes binding incubation and 10 minutes elution at 37°C. All other steps were performed following the manufacturer’s protocol. Libraries were sequenced using the best available flow cells at the time (FLO-MAP103 R7). Libraries were topped-up after 12 hours, and sequencing performed for 24 hours in total. Data were base-called in real-time via Metrichor, using the best available workflow at the time (ONT, UK).

WGS data processing and analysis

Following confirmation of species identification, Illumina short-reads were mapped to *E. coli* references: *Escherichia coli* CFT073 (GenBank accession number AE014075.1) for the complete dataset of *E. coli* sequences, and also an ST216 *E. coli*-specific reference generated from this study - namely the H151860951 chromosome - derived from PacBio sequencing (see below). Mapping and variant calling, and associated flags and filters, are set up as part of an in-house bioinformatics pipeline, as in².

The phylogeny for the *E. coli* isolates was reconstructed with IQTree³, using a GTR+G model and a maximum parsimony starting tree. The consensus fasta sequence outputs from the mapping pipeline were used as input. Subsequent correction for recombination performed by running ClonalFrameML with default parameters⁴.

Short-read data were also assembled using SPAdes (v3.6)⁵. All isolates had assembly size < 6.5 Mb. Of the 292 sequenced isolates, n=24 isolates were not included in analysis as they represented replicate isolates from single sample cultures. *In silico* multi-locus sequence-typing (MLST) was performed using the MLST database for *E. coli* downloaded from <http://mlst.warwick.ac.uk/mlst/dbs/Ecoli>, using BLASTn and a 100% sequence length/identity match as the threshold for calling alleles. Plasmid Inc typing was performed using the PlasmidFinder database⁶ with 80% sequence identity and 90% coverage as

thresholds, and Tn4401 typing by BLASTn against a catalogue of known isoforms. Flanking sequences of the Tn4401 elements were identified by obtaining the 5bp sequences at the boundary where matches against the reference Tn4401 were identified. Resistance genotyping was performed using resistType (https://github.com/hangphan/resistType_docker), a BLASTn/local-reassembly-based approach for identifying presence/absence of known chromosomal and acquired resistance mechanisms.

For plasmid typing, we downloaded all publically available complete plasmids from NCBI (query term: plasmids AND Enterobacteriaceae AND complete sequence), de-duplicated them, and extracted any plasmids carrying *bla*_{KPC}. In addition, fully closed KPC-plasmid sequences from a global KPC study⁷ and KPC-plasmids from the two KPC-*E. coli* isolates in this study sequenced using long-read methods were added to our reference KPC plasmid database (n=194 KPC plasmids), which were then de-duplicated by pairwise comparisons to remove those that shared >95% sequence similarity over >90% of any other plasmid sequence's length, leaving 134 distinct KPC-plasmids in the final database. KPC plasmids in the final database were then used as *in silico* probes to identify whether similar *bla*_{KPC}-carrying plasmid structures might be present in the study *E. coli* isolates, again using BLASTn. An isolate was assigned as putatively harbouring a candidate reference KPC-plasmid if a total of >75% of the reference plasmid could be aligned with contigs derived from the isolate's assembly with the following threshold: >90% of the combined contig lengths matching the reference plasmid with >95% sequence identity.

From MinION sequencing data, 2D reads passing quality control were extracted using poretools⁸. HybridSPAdes (v3.6)⁵ combining short-read Illumina data and long-read MinION sequencing data, and Canu (v1.4)⁹ assembler + pilon (v1.18)¹⁰, an assembly "polisher", were used to generate hybrid assemblies. The assembly-polishing step was used because MinION Canu assemblies are prone to nucleotide-level error, which can be corrected using short-read Illumina data (with very low error rates). Dnadiff and mummerplot¹¹ were used to identify discrepancies between assemblies from these two assembly pipelines. Assembly integrity was checked using REAPR¹², and any discrepancies were manually inspected and amended to obtain fully circularised *bla*_{KPC} plasmid structures where possible. We defined a plasmid as circularised if it had >100bp overlapping ends with 100% sequence identity for hybridSPAdes assemblies, and >1kbp overlapping ends at >99% sequence identity for Canu+pilon assemblies.

PacBio sequence assembly results

The PacBio assembly of the outbreak isolate H151860951 had two closed contigs, namely a chromosome of 4.73Mb, and a circularized plasmid sequence, pKPC-CAD1 (200kb). The chromosomal sequence was used as the reference for mapping short-read Illumina data from the other ST216 isolates to generate consensus fasta sequences (including variant sites) for the ST216 *E. coli*-specific phylogenetic analysis.

MinION sequencing data – summary statistics

The MinION sequencing run for the 2012 ST216 isolate resulted in ~30x yield, with a mean read length of 7.17kb, and a maximum read length reaching 40kb. The consensus, manually checked set of three contigs generated from Canu/pilon and hybridSPAdes assemblies included a 4.73Mb chromosomal sequence, one circularised *bla*_{KPC}-plasmid pKPC-CAD2 (307kb) and another circularised plasmid pCAD3 (153kb), highly similar to pKPC-CAD1 (see main text).

Table S1. Summary of standard infection prevention and control interventions implemented Trust-wide and enhanced measures implemented in response to the Manchester Heart Centre KPC-*Escherichia coli* outbreak.

Infection control measure	From July 2014 through to outbreak identification in April 2015	In response to outbreak
Patient CRE screening	<ul style="list-style-type: none"> • Routine screening* of high-risk** patients on admission and weekly • Screening of all patients on affected wards following identification of any CRE-positive patients (either colonised or clinical cases) • Twice weekly screening* if ward closed (see below) until no new cases identified 	<ul style="list-style-type: none"> • Routine twice weekly screening* of inpatients until no new cases identified from two consecutive screens taken 48 hours apart
Bed occupancy	<ul style="list-style-type: none"> • No intervention 	<ul style="list-style-type: none"> • W3 reduced from 28 in-patient beds to 10 day cases only (August 2015) • W4 reduced 28 in-patient beds to 12 in-patient beds only (August 2015)
Access to wards	<ul style="list-style-type: none"> • No intervention 	<ul style="list-style-type: none"> • Access to wards restricted to essential medical staff • Connecting doors between W3 and W4

Infection control measure	From July 2014 through to outbreak identification in April 2015	In response to outbreak
Ward closures	<ul style="list-style-type: none"> • Wards closed if more than two CRE-positive patients identified on routine screening* • Re-opened only if all CRE-negative patients remained negative on two screens* 48 hours apart and CRE-positive patients could be isolated/cohorted or moved to dedicated CRE cohort wards 	<ul style="list-style-type: none"> • Additionally, W3/W4 completely closed to allow plumbing replacement (Sept 2015)
Hand hygiene	<ul style="list-style-type: none"> • Standard WHO guidelines 	<ul style="list-style-type: none"> • Staff monitors at ward entrances to check adherence to guidelines • Patients asked to comment anonymously on hand hygiene practices
Contact precautions	<ul style="list-style-type: none"> • Standard precautions 	<ul style="list-style-type: none"> • Long-sleeved disposable single use gowns to care for CRE-positive patients
Review of facilities	<ul style="list-style-type: none"> • Routine 	<ul style="list-style-type: none"> • Weekly review of the MHC by the Head of Nursing and Senior Domestic Facilities Manager to ensure environmental cleaning standards maintained • Integrity of the mattresses and pillows inspected and damaged/breached items replaced
Environmental cleaning	<ul style="list-style-type: none"> • Hypochlorite 1000 ppm 	<ul style="list-style-type: none"> • Hypochlorite 1000 ppm - frequency of application increased from twice daily to three times daily for high risk surfaces i.e. patient toilets/locker tops and bed tables • Hydrogen peroxide vapour used to decontaminate during outbreaks
Environmental sampling and measures	<ul style="list-style-type: none"> • No intervention 	<ul style="list-style-type: none"> • Intermittent environmental sampling with 13 sampling events on W3/W4 and nine on non-W3/W4 wards, including sinks, drains,

Infection control measure	From July 2014 through to outbreak identification in April 2015	In response to outbreak
Antimicrobial Stewardship	<ul style="list-style-type: none"> • Antimicrobial policy emphasising alternatives to carbapenem usage • Written approval for meropenem use required from September 2014 	<ul style="list-style-type: none"> • toilets/hoppers and high-touch surfaces • Decontamination of the sink outlets and replacement of the sink trap as part of pre-ward opening protocols • Bottled water distributed to patients when KPC-EC isolated from W3/W4 kitchen sinks/drains • Additional CMFT-wide daily clinical review of patients prescribed carbapenems from May 2015
Transfer protocols	<ul style="list-style-type: none"> • All patients treated at MHC were subsequently managed as ‘suspected’ cases of CRE on return to referring hospitals, requiring isolation until three consecutive screening samples (taken 48 hours apart) tested negative 	<ul style="list-style-type: none"> • In consultation with the regional Strategic Cardiac Clinical Network Group and PHE, guidance was adapted to allow patient flow and urgent treatment, whilst maintaining appropriate measures to detect and manage any CRE-positive patients

* ‘screen’=rectal screening for CRE

** high-risk defined as previously CRE positive or history of hospitalization abroad or in UK hospital with a known CRE problem in last 12 months.

Table S2. CRE screening program in Central Manchester Foundation NHS Trust (CMFT)

Period characteristic	Early P1	Late P1	Early P2	Late P2	P3	P4
Start	18/01/2010	30/06/2014	01/09/2014	01/01/2015	24/09/2015	15/01/2016
End	29/06/2014	31/08/2014	31/12/2014	23/09/2015	14/01/2016	31/12/2016
Period duration in days	1,624	63	122	266	113	352
Number of patients admitted	173,944	15,481	27,991	53,117	26,912	65,866
Number of admissions	771,572	32,232	64,843	145,437	60,387	193,125
Number of overnight stays	1,816,348	75,349	152,951	321,464	145,906	426,117
Number of CRE screens performed by culture or molecular tests	36,349	4,399	15,281	46,608	18,876	63,026
Number of CRE screens positive by culture or molecular tests	1,769	218	425	827	134	785
Number of patients CRE screened by culture or molecular tests	16,141	2,102	6,289	16,803	8,119	20,869
Number of patients identified as CRE positive by culture or molecular tests	1,367	152	265	502	112	426
Number of CRE screens performed by culture	36,224	4,059	8,820	21	0	37
Number of CRE screens positive by culture	1,768	211	272	2	0	0
Number of patients CRE screened by culture	16,089	1,964	3,861	20	0	34
Number of patients identified as CRE positive by culture	1,366	148	181	1	0	0
Number of CRE screens performed by molecular tests	125	340	6,461	46,587	18,876	62,989
Number of CRE screens positive by molecular tests	1	7	153	825	134	785
Number of patients CRE screened By molecular tests	107	299	4,087	16,803	8,119	20,852
Number of patients identified as CRE positive by molecular tests	1	6	104	501	112	426
Number of molecular screens with mechanisms available	0	6	128	643	116	689
Number of molecular screens with mechanisms not available	1	1	25	182	18	96
Number of molecular screens with mixed CR mechanisms detected	0	0	4	6	3	23
Number of molecular screens with KPC mechanisms detected	0	6	124	617	110	620
Number of molecular screens with OXA mechanisms detected	0	0	1	10	2	50
Number of molecular screens with VIM mechanisms detected	0	0	2	3	0	0
Number of molecular screens with NDM mechanisms detected	0	0	4	3	3	17
Number of molecular screens with IMP mechanisms detected	0	0	0	0	0	4

Table S3. Details of sequenced isolates

All isolates contained *bla_{KPC-2}*, except for isolate HI134340590, which contained *bla_{KPC-3}* and isolates without *bla_{KPC}* (N/A in “Tn4401 and flanking sequences” column). “CR” = carbapenem-resistant.

Isolate name	Biosample accession	Date of collection	Assembly size	ST	Tn4401 and flanking sequences	Inclusion type
HI131800734	SAMNN06630144	16/04/13	4809457	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
14.2002125	SAMNN06630091	13/10/14	5148953	409	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
14.2002257	SAMNN06630078	14/10/14	5151133	409	Tn4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
HI151860924	SAMNN06630197	24/02/15	4815167	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
HI151860926	SAMNN06630246	03/03/15	4820409	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
2006447 C10	SAMNN06630333	10/03/15	5146955	401	Tn4401-2 9907 AGTTG AGTTG	environmental
2006447 D1	SAMNN06630332	10/03/15	5129374	401	Tn4401-2 9907 AGTTG AGTTG	environmental
2006447 D2	SAMNN06630330	10/03/15	4883501	216	Tn4401-2 9907 AGTTG AGTTG	environmental
2006447 D3	SAMNN06630339	10/03/15	5148064	401	Tn4401-2 9907 AGTTG AGTTG	environmental
2006447 D4	SAMNN06630336	10/03/15	5168378	401	Tn4401-2 9907 AGTTG AGTTG	environmental
HI152240903	SAMNN06630134	22/03/15	4823394	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
HI151860928	SAMNN06630199	30/03/15	4824059	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
HI151860929	SAMNN06630203	30/03/15	4829463	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
HI151660768	SAMNN06630278	02/04/15	4822457	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
HI151660769	SAMNN06630188	06/04/15	4818653	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
HI151860930	SAMNN06630276	06/04/15	4823283	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2211041.V	SAMNN06630286	06/04/15	5111591	155	Tn4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
HI151640734	SAMNN06630160	08/04/15	4882136	216	Tn4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
HI151640735	SAMNN06630146	08/04/15	4818943	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
HI151660767	SAMNN06630119	08/04/15	4821554	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
HI151640736	SAMNN06630150	09/04/15	4820345	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
HI151980637	SAMNN06630310	12/04/15	4817532	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
HI151660766	SAMNN06630129	12/04/15	4817125	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
HI151660773	SAMNN06630204	12/04/15	4818725	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
HI151660775	SAMNN06630163	12/04/15	4817582	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>

Isolate name	Biosample accession	Date of collection	Assembly size	ST	Th4401 and flanking sequences	Inclusion type
H151860931	SAMN006630248	16/04/15	4822934	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H151860932	SAMN006630314	18/04/15	4823107	216	Th4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
H151980640	SAMN006630137	20/04/15	4819618	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H151860951	SAMN006630284	22/04/15	4813332	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H151860952	SAMN006630140	22/04/15	4823362	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H151860953	SAMN006630157	22/04/15	4820390	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H151980642	SAMN006630279	22/04/15	4827407	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H151920660	SAMN006630106	26/04/15	5010629	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2213525.K	SAMN006630239	26/04/15	4799667	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2213527.B	SAMN006630177	26/04/15	4803097	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2213542.P	SAMN006630224	26/04/15	4805104	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H151920661	SAMN006630153	26/04/15	4818731	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2213935.L	SAMN006630192	30/04/15	4805688	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2214340.J	SAMN006630300	03/05/15	4818549	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2214923.M	SAMN006630295	07/05/15	4803294	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H152240898	SAMN006630334	14/05/15	4808460	216	Th4401-2 9907 AGTTG AGTTG	environmental
H152160660	SAMN006630331	14/05/15	4808323	216	Th4401-2 9907 AGTTG AGTTG	environmental
H152160661	SAMN006630329	14/05/15	4819335	216	Th4401-2 9907 AGTTG AGTTG	environmental
H152240897	SAMN006630335	14/05/15	4819739	216	Th4401-2 9907 AGTTG AGTTG	environmental
H152180644	SAMN006630266	18/05/15	4820908	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H152240990	SAMN006630214	18/05/15	4811224	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H152240899	SAMN006630337	18/05/15	4824511	216	Th4401-2 9907 AGTTG AGTTG	environmental
H152180645	SAMN006630113	18/05/15	4823223	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H152180646	SAMN006630099	18/05/15	4820119	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H152180648	SAMN006630215	18/05/15	4812778	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H152180649	SAMN006630079	18/05/15	4817641	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
2012194		19/05/15	4786345	216	Th4401-2 9907 AGTTG AGTTG	environmental
H152240976	SAMN006630328	19/05/15	4819010	216	Th4401-2 9907 AGTTG AGTTG	environmental
M.15.2216647.L	SAMN006630249	20/05/15	4803560	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>

Isolate name	Biosample accession	Date of collection	Assembly size	ST	Th4401 and flanking sequences	Inclusion type
H152280659	SAMN006630067	22/05/15	4820927	216	Th4401-2 9907 AGTTG AGTTG	first CR-E. coli
M.15.2220743.G	SAMN006630171	18/06/15	5609219	401	Th4401-2 9907 AGTTG AGTTG	first CR-E. coli
M.15.2223010.J	SAMN006630087	02/07/15	4796799	216	Th4401-2 9907 AGTTG AGTTG	first CR-E. coli
M.15.2223938.G	SAMN006630127	09/07/15	4804942	216	Th4401-2 9907 AGTTG AGTTG	first CR-E. coli
M.15.2224369.W	SAMN006630262	13/07/15	4947956	131	Th4401-2 9907 ATTGA ATTGA	first CR-E. coli
M.15.2225907.Q	SAMN006630240	30/07/15	4819046	216	Th4401-2 9907 AGTTG AGTTG	first CR-E. coli
M.15.2226214.N	SAMN006630062	03/08/15	4823451	216	Th4401-2 9907 AGTTG AGTTG	first CR-E. coli
M.15.2226219.T	SAMN006630083	03/08/15	4811719	216	Th4401-2 9907 AGTTG AGTTG	first CR-E. coli
M.15.2226227.D	SAMN006630252	03/08/15	4817642	216	Th4401-2 9907 AGTTG AGTTG	first CR-E. coli
M.15.2226216.Y	SAMN006630244	03/08/15	4823972	216	Th4401-2 9907 AGTTG AGTTG	first CR-E. coli
M.15.2226203.C	SAMN006630175	03/08/15	4816390	216	Th4401-2 9907 AGTTG AGTTG	first CR-E. coli
M.15.2226210.B	SAMN006630082	03/08/15	4815294	216	Th4401-2 9907 AGTTG AGTTG	first CR-E. coli
M.15.2226505.T	SAMN006630237	06/08/15	4830984	216	Th4401-2 9907 AGTTG AGTTG	first CR-E. coli
M.15.2226517.H	SAMN006630201	06/08/15	4807577	216	Th4401-2 9907 AGTTG AGTTG	first CR-E. coli
M.15.2226715.S.1	SAMN006630236	08/08/15	4802469	216	Th4401-2 9907 AGTTG AGTTG	first CR-E. coli
M.15.2226818.E	SAMN006630238	10/08/15	4806214	216	Th4401-2 9907 AGTTG AGTTG	first CR-E. coli
M.15.2021426.H	SAMN006630207	13/08/15	4822715	216	Th4401-2 9907 AGTTG AGTTG	first CR-E. coli
M.15.2021947.Q	SAMN006630307	14/08/15	5026360	216	Th4401-2 9907 AGTTG AGTTG	first CR-E. coli
M.15.2022950.F	SAMN006630275	24/08/15	4815391	216	Th4401-2 9907 AGTTG AGTTG	first CR-E. coli
2023751_3		07/09/15	4785030	216	Th4401-2 9907 AGTTG AGTTG	environmental
1950044_2	SAMN006630285	08/02/16	5417856	216	Th4401-2 9907 AGTTG AGTTG	environmental
1950343_2_1	SAMN006630338	21/03/16	4826428	216	Th4401-2 9907 AGTTG AGTTG	environmental
16.2020790_10.2	SAMN006630138	31/05/16	5037566	410	Th4401-2 9907 ATTGA ATTGA	sequential CR-E. coli
16.2020416_1	SAMN006630128	31/05/16	5099461	401	Th4401-2 9907 AGTTG AGTTG	sequential CR-E. coli
2015589_1_1	SAMN006630169	31/05/16	5104052	401	Th4401-2 9907 AGTTG AGTTG	unlinked*
1951413_5	SAMN006630104	02/08/16	4895692	10	Th4401-2 9907 ATTGA ATTGA	environmental
H102400241	SAMN006630306	07/06/10	5284489	372	Th4401-2 9907 ATTGA ATTGA	unlinked*
H104320602	SAMN006630118	16/10/10	4932859	410	Th4401-2 9907 ATTGA ATTGA	first CR-E. coli
H104500514	SAMN006630290	26/10/10	4976189	635	Th4401-2 9907 ATTGA ATTGA	first CR-E. coli

Isolate name	Biosample accession	Date of collection	Assembly size	ST	Th4401 and flanking sequences	Inclusion type
H1104400237	SAMN006630081	26/10/10	5106842	1276	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
H104620602	SAMN006630223	08/11/10	5361684	5667	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
H110300564	SAMN006630211	15/01/11	5198220	131	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
H111000292	SAMN006630311	28/02/11	4924357	410	Th4401-2 9907 ATTGA ATTGA	unlinked*
H112040301	SAMN006630218	09/05/11	5074937	131	Th4401-2 9907 ATTGA ATTGA	unlinked*
H112260216	SAMN006630283	23/05/11	5279722	New	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
H112400221	SAMN006630147	31/05/11	5075116	635	Th4401-2 9907 ATTGA ATTGA	unlinked*
H112480151	SAMN006630257	07/06/11	5384159	131	Th4401-2 9907 ATTGA ATTGA	unlinked*
H112480153	SAMN006630265	08/06/11	4933462	93	Th4401-2 9907 ATTGA ATTGA	unlinked*
H112600372	SAMN006630102	17/06/11	5213421	141	Th4401-2 9907 ATTGA ATTGA	unlinked*
H112800328	SAMN006630241	29/06/11	5065213	90	Th4401-2 9907 ATTGA ATTGA	unlinked*
H113200478	SAMN006630114	01/08/11	4832410	1415	Th4401-2 9907 ATTGA ATTGA	unlinked*
H113540400	SAMN006630098	19/08/11	4913676	410	Th4401-2 9907 ATTGA ATTGA	unlinked*
H114720537	SAMN006630206	10/11/11	4913379	410	Th4401-2 9907 ATTGA ATTGA	unlinked*
H114800132	SAMN006630221	16/11/11	4926112	410	Th4401-2 9907 ATTGA ATTGA	unlinked*
H125200698	SAMN006630096	03/12/12	4806563	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H133600452	SAMN006630111	21/08/13	5062481	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H133740327	SAMN006630136	31/08/13	5060917	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H134540677	SAMN006630191	28/10/13	5491794	404	Th4401-2 9907 ATTGA ATTGA	sequential CR- <i>E. coli</i>
13.162628	SAMN006630100	30/12/13	4761794	216	Th4401-2 9907 AGTTG AGTTG	unlinked*
14.1611529	SAMN006630181	25/01/14	4994695	607	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
14.1611427	SAMN006630183	27/01/14	5292308	327	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
14.1313153	SAMN006630268	08/02/14	4821833	2802	Th4401-2 9907 ATTGA ATTGA	sequential CR- <i>E. coli</i>
14.131401	SAMN006630075	10/02/14	4921075	216	Th4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
14.1612399	SAMN006630172	11/02/14	5251548	607	Th4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
14.1612942	SAMN006630228	18/02/14	6036937	216	Th4401-2 9907 AGTTG	first CR- <i>E. coli</i>
14.1613028	SAMN006630148	20/02/14	5571662	404	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
14.1613223	SAMN006630125	24/02/14	4827590	216	Th4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
14.1613784	SAMN006630184	05/03/14	5207157	401	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>

Isolate name	Biosample accession	Date of collection	Assembly size	ST	Th4401 and flanking sequences	Inclusion type
H141520767	SAMN006630291	25/03/14	4825766	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
14.1623262	SAMN006630267	19/08/14	4776033	10	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
14.1624814	SAMN006630063	05/09/14	5842476	607	Th4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
14.1624931	SAMN006630117	08/09/14	5347426	401	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
14.2000662	SAMN006630077	29/09/14	5275688	607	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
14.1392673	SAMN006630126	08/10/14	5422601	10	Th4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
14.2002178	SAMN006630154	13/10/14	4897648	410	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
14.2100321	SAMN006630281	15/10/14	4998010	607	Th4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
14.2100451	SAMN006630255	20/10/14	4820675	2802	Th4401-2 9907 ATTGA ATTGA	sequential CR- <i>E. coli</i>
14.2100566	SAMN006630131	22/10/14	5251841	607	Th4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
14.2003293	SAMN006630194	22/10/14	5145418	409	Th4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
14.2100664	SAMN006630135	24/10/14	5012098	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
14.2101104	SAMN006630089	27/10/14	5424699	404	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
14.100067	SAMN006630176	31/10/14	4776139	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
14.2100949	SAMN006630151	02/11/14	4813248	216	Th4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
14.2100962	SAMN006630312	03/11/14	4819394	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
14.2102106	SAMN006630112	13/11/14	5318945	131	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
14.2005586	SAMN006630152	17/11/14	5169210	410	Th4401-2 9907 ATTGA ATTGA	sequential CR- <i>E. coli</i>
14.2005903	SAMN006630189	19/11/14	4984675	10	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
14.2007212_1	SAMN006630156	02/12/14	5612147	940	Th4401a 46 9907 ATTGA ATTGA	sequential CR- <i>E. coli</i>
14.2104472	SAMN006630254	24/12/14	5175601	131	Th4401-2 9907 ATTGA ATTGA	sequential CR- <i>E. coli</i>
M.15.2200338.F	SAMN006630132	06/01/15	4932694	131	Th4401-2 9907 ATTGA ATTGA	sequential CR- <i>E. coli</i>
M.15.2000946.G	SAMN006630213	07/01/15	5430676	141	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
M.15.2002069.C	SAMN006630251	19/01/15	5024201	349	Th4401-2 9907 ATTGA ATTGA	sequential CR- <i>E. coli</i>
M.15.2201773.W	SAMN006630292	21/01/15	5074788	401	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2202865.W	SAMN006630260	31/01/15	5228363	131	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
M.15.2003439.Q	SAMN006630071	04/02/15	5522560	929	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2203694.Y	SAMN006630270	07/02/15	5116751	362	Th4401-2 9907 ATTGA ATTGA	sequential CR- <i>E. coli</i>
M.15.2004097.J	SAMN006630122	11/02/15	5026242	2329	Th4401-2 9907 ATTGA ATTGA	sequential CR- <i>E. coli</i>

Isolate name	Biosample accession	Date of collection	Assembly size	ST	Th4401 and flanking sequences	Inclusion type
M.15.2204295.L	SAMN006630282	13/02/15	4789014	216	Th4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
M.15.2205319.F	SAMN006630247	23/02/15	5165469	1193	Th4401-2 9907 ATTGA ATTGA	sequential CR- <i>E. coli</i>
M.15.2005197.G	SAMN006630085	24/02/15	5330929	404	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
M.15.2205516.N	SAMN006630195	26/02/15	4993887	410	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
H151080744	SAMN006630170	27/02/15	4852812	216	Th4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
M.15.2205863.W	SAMN006630124	01/03/15	4865005	410	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
M.15.2005802.K	SAMN006630293	02/03/15	5744615	69	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2210170.B	SAMN006630234	29/03/15	5246130	131	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
H151660772	SAMN006630269	09/04/15	4816273	216	Th4401-2 9907 AGTTG AGTTG	unlinked*
H151980638	SAMN006630065	13/04/15	4825490	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H151980639	SAMN006630287	14/04/15	5108107	216	Th4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
M.15.2009402.W	SAMN006630178	15/04/15	4846267	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2212636.B	SAMN006630187	19/04/15	4847313	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H151980641	SAMN006630216	20/04/15	4815052	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H151860950	SAMN006630168	21/04/15	4906091	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2213598.C	SAMN006630227	27/04/15	5417869	405	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
M.15.2011001.N	SAMN006630222	03/05/15	4798275	216	Th4401-2 9907 AGTTG AGTTG	unlinked*
M.15.2011460.S	SAMN006630235;	09/05/15	4957116	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2215257.H	SAMN006630298	09/05/15	5091378	401	Th4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
M.15.2011745.H	SAMN006630205	13/05/15	4860185	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2217771.W	SAMN006630076	28/05/15	4806684	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2219817.F	SAMN006630120	11/06/15	5335822	1485	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
M.15.2014751.K	SAMN006630179	17/06/15	4824926	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2222633.W	SAMN006630115	29/06/15	4735588	1415	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
M.15.2015974.E	SAMN006630084	01/07/15	4761007	6178	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
M.15.2223601.S	SAMN006630108	05/07/15	5082239	New	Th4401-2 9907 ATTGA ATTGA	unlinked*
M.15.2224004.D	SAMN006630174	08/07/15	5307466	127	Th4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
M.15.2017032.N	SAMN006630165	13/07/15	4887764	216	Th4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2017831.J	SAMN006630101	19/07/15	5236202	69	Th4401a 44 9907 ATTGA ATTGA	first CR- <i>E. coli</i>

Isolate name	Biosample accession	Date of collection	Assembly size	ST	Tn4401 and flanking sequences	Inclusion type
M.15.2225980.V	SAMN006630185	30/07/15	5078123	401	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2019768.J	SAMN006630230	31/07/15	4851785	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2021028.K	SAMN006630245	10/08/15	5432833	New	Tn4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
M.15.2021381.S	SAMN006630166	12/08/15	5410649	401	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2021399.E	SAMN006630173	12/08/15	5410627	401	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2021407.Y	SAMN006630304	12/08/15	5397819	401	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2023374.C	SAMN006630308	27/08/15	4859194	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2023801.P	SAMN006630208	01/09/15	5415136	401	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2024339.H	SAMN006630086	05/09/15	5154892	401	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2024771.F	SAMN006630229	09/09/15	5304446	69	Tn4401-2 9907 Tn4401-2 9907	first CR- <i>E. coli</i>
2213305_1	SAMN006630105	11/07/16	5244403	405	Tn4401-2 9907 TAATA TAATA	unlinked*
2213346_3	SAMN006630133	11/07/16	5236889	69	Tn4401-2 9907 TAATA CTAAT	unlinked*
H104320604	SAMN006630190	25/10/10	5134397	635	Tn4401-2 9907 ATTGA ATTGA	sequential CR- <i>E. coli</i>
H112600370	SAMN006630261	16/06/11	5406831	69	Tn4401-2 9907 ATTGA ATTGA	unlinked*
H112800327	SAMN006630233	04/07/11	5266047	69	Tn4401-2 9907 ATTGA ATTGA	unlinked*
H124200644	SAMN006630164	04/10/12	5038560	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H124200645	SAMN006630162	04/10/12	5059900	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H124200646	SAMN006630258	05/10/12	5067523	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H124200647	SAMN006630196	05/10/12	5063888	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H124300748	SAMN006630143	12/10/12	5065665	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H124300749	SAMN006630289	12/10/12	5120383	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H124400498	SAMN006630110	19/10/12	5175740	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H125200693	SAMN006630182	25/10/12	5221544	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H124600287	SAMN006630094	02/11/12	5071807	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H124600288	SAMN006630080	02/11/12	5103443	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H124900514	SAMN006630210	04/11/12	5061282	216	Tn4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
H124900518	SAMN006630209	12/11/12	5128374	216	Tn4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
H124900520	SAMN006630141	16/11/12	5171110	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H124900524	SAMN00663027	16/11/12	5071256	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>

Isolate name	Biosample accession	Date of collection	Assembly size	ST	Tn4401 and flanking sequences	Inclusion type
H124900519	SAMN006630243	19/11/12	5071651	216	Tn4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
H125200695	SAMN006630217	23/11/12	5078393	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H125000203	SAMN006630220	30/11/12	4919261	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H125200696	SAMN006630068	30/11/12	5069139	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H125100459	SAMN006630088	04/12/12	4806898	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H125200694	SAMN006630130	07/12/12	4806271	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H125200499	SAMN006630225	14/12/12	5050864	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
14.1611868	SAMN006630070	01/02/14	4816945	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
14.161354	SAMN006630097	02/03/14	5365914	10	Tn4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
14.2003512	SAMN006630288	25/10/14	4918435	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
14.2101329	SAMN006630158	28/10/14	4875785	216	Tn4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
H152280657	SAMN006630193	20/05/15	4812662	216	Tn4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
M.15.2012485.X	SAMN006630250	20/05/15	5092389	409	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2226022.T	SAMN006630064	30/07/15	4859562	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2020367.Z	SAMN006630253	06/08/15	5091847	401	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M.15.2023155.M	SAMN006630301	26/08/15	5379068	401	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
2213314_1	SAMN006630297	10/07/16	5251341	405	Tn4401-2 9907 TAATA TAATA	unlinked*
2213347_4	SAMN006630299	11/07/16	4980778	405	Tn4401-2 9907 TAATA TAATA	unlinked*
2213348_5	SAMN006630231	11/07/16	5244810	69	Tn4401-2 9907 CTAAT TAATA	unlinked*
H111860374	SAMN006630313	28/04/11	5167532	131	Tn4401-2 9907 ATTGA ATTGA	regional study
H112980482	SAMN006630271	16/07/11	5434476	404	Tn4401-2 9907	regional study
H113520207	SAMN006630242	19/08/11	5303538	131	Tn4401-2 9907 ATTGA ATTGA	regional study
H115160637	SAMN006630066	13/12/11	5345271	Nov el ST	Tn4401-2_9907_ATTGA_ATTGA	regional study
H120740613	SAMN006630296	07/02/12	5394465	95	Tn4401-2 9907 ATTGA ATTGA	regional study
H121020572	SAMN006630167	20/02/12	4908733	410	Tn4401-2 9907 ATTGA ATTGA	regional study
H121580296	SAMN006630202	06/04/12	4973409	58	Tn4401-2 9907 ATTGA ATTGA	regional study
H121940564	SAMN006630095	29/04/12	5391511	95	Tn4401-2 9907 ATTGA ATTGA	regional study

Isolate name	Biosample accession	Date of collection	Assembly size	ST	Tn4401 and flanking sequences	Inclusion type
H122480255	SAMN006630149	31/05/12	4814091	216	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H124260425	SAMN006630273	04/07/12	4924646	410	Tn4401-2 9907 ATTGA ATTGA	unlinked*
H123180719	SAMN006630277	25/07/12	5161728	3858	Tn4401-2 9907 ATTGA ATTGA	regional study
H124580518	SAMN006630159	19/09/12	5135484	127	Tn4401-2 9907 ATTGA ATTGA	regional study
H124820234	SAMN006630072	20/11/12	5079519	155	Tn4401-2 9907 ATTGA ATTGA	unlinked*
H130780386	SAMN006630305	05/02/13	5286554	7075	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
H132500545	SAMN006630294	10/06/13	5290758	95	Tn4401-2 9907 ATTGA ATTGA	regional study
H132800276	SAMN006630309	29/06/13	5352520	12	Tn4401-2 9907 ATTGA ATTGA	regional study
H134340590	SAMN006630142	18/10/13	5148433	58	Tn4401a 14 9907 ATTGA ATTGA	unlinked*
H134820403	SAMN006630198	22/11/13	5607670	73	Tn4401-2 9907 AGTTG AGTTG	unlinked*
H140520551	SAMN006630186	23/01/14	5184670	131	Tn4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
H140780383	SAMN006630264	05/02/14	5396496	131	Tn4401-2 9907 ATTGA ATTGA	sequential CR- <i>E. coli</i>
H142140855	SAMN006630259	16/05/14	5301045	95	Tn4401-2 9907 ATTGA ATTGA	regional study
H142120710	SAMN006630121	20/05/14	4761419	216	Tn4401-2 9907 AGTTG AGTTG	regional study
H143220568	SAMN006630302	09/07/14	5416750	372	Tn4401-2 9907 AGTTG AGTTG	regional study
H143120731	SAMN006630226	21/07/14	5246922	131	Tn4401-2 9907 ATTGA ATTGA	regional study
H143420799	SAMN006630232	11/08/14	5159033	1193	Tn4401-2 9907 ATTGA ATATA	regional study
H143520204	SAMN006630116	15/08/14	5150844	1193	Tn4401-2 9907 ATTGA ATATA	regional study
H143540714	SAMN006630212	18/08/14	5206209	131	Tn4401-2 9907 ATTGA ATTGA	regional study
14.2002212	SAMN006630200	13/10/14	5014903	216	Tn4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
H144620795	SAMN006630107	04/11/14	5188510	2227	Tn4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
M15.0367554.M	SAMN006630145	09/02/15	5220706	69	Tn4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
M15.0368093.H 1	SAMN006630090	25/02/15	4871506	91	Tn4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
M15.0368991.J	SAMN006630315	26/03/15	5230262	69	Tn4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
M.15.2221646.V	SAMN006630139	24/06/15	5234354	38	Tn4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
M15.0372899.D	SAMN006630272	14/07/15	4813540	216	Tn4401-2 9907 AGTTG AGTTG	sequential CR- <i>E. coli</i>
M15.0374051.H	SAMN006630263	11/08/15	5286177	131	Tn4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
M15.0374971.D	SAMN006630180	06/09/15	5170172	New	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M15.0375247.T 2	SAMN006630155	14/09/15	5103068	131	Tn4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>

Isolate name	Biosample accession	Date of collection	Assembly size	ST	Tn4401 and flanking sequences	Inclusion type
M15.0376073.H	SAMNN06630256	03/10/15	5166443	New	Tn4401-2 9907 AGTTG AGTTG	first CR- <i>E. coli</i>
M15.0376392.Z	SAMNN06630123	10/10/15	5447255	131	Tn4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
M15.0376606.K	SAMNN06630103	14/10/15	5114114	131	Tn4401-2 9907 ATTGA ATTGA	sequential CR- <i>E. coli</i>
M15.0376751.V	SAMNN06630219	19/10/15	5162213	20	Tn4401-2 9907 ATTGA ATTGA	first CR- <i>E. coli</i>
H134880341	SAMNN06630161	19/11/13	4793018	216	N/A	no KPC
M.15.2009078.J	SAMNN06630092	09/04/15	5162664	38	N/A	no KPC
14.1624936	SAMNN06630109	08/09/14	5381695	69	N/A	no KPC
M.15.2000950.G	SAMNN06630280	07/01/15	5245448	38	N/A	no KPC
M.15.2001889.R	SAMNN06630069	17/01/15	3653868	New	N/A	no KPC
M.15.2220818.X	SAMNN06630073	17/06/15	4894406	156	N/A	no KPC
M.15.2221656.H	SAMNN06630303	21/06/15	5220254	1882	N/A	no KPC
M.15.2223686.K	SAMNN06630074	06/07/15	5208654	38	N/A	no KPC
M.15.2021393.J	SAMNN06630093	12/08/15	4825029	3541	N/A	no KPC

* unlinked epidemiological information

Supplementary Figures

Figure S1A. Schematic layout of Manchester Heart Centre (MHC wards) within Manchester Royal Infirmary. Wards associated with the MHC are shown in i

A.

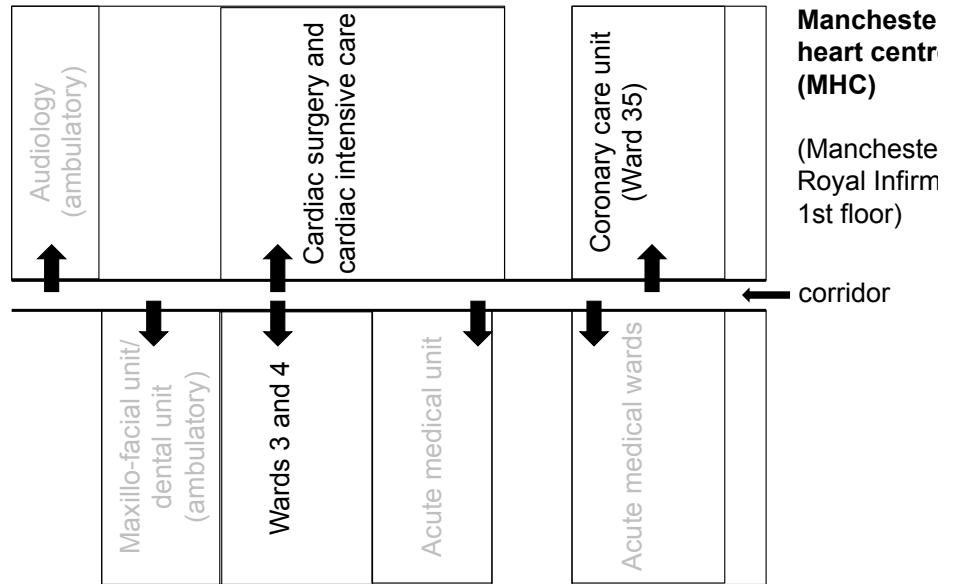


Figure S1B. Schematic layout of Cardiac wards 3 and 4. Both ward dirty utility were colonised with CRE on sampling of the ward after patient relocation and plumbing intervention, and just prior to re-opening of the wards.

B.

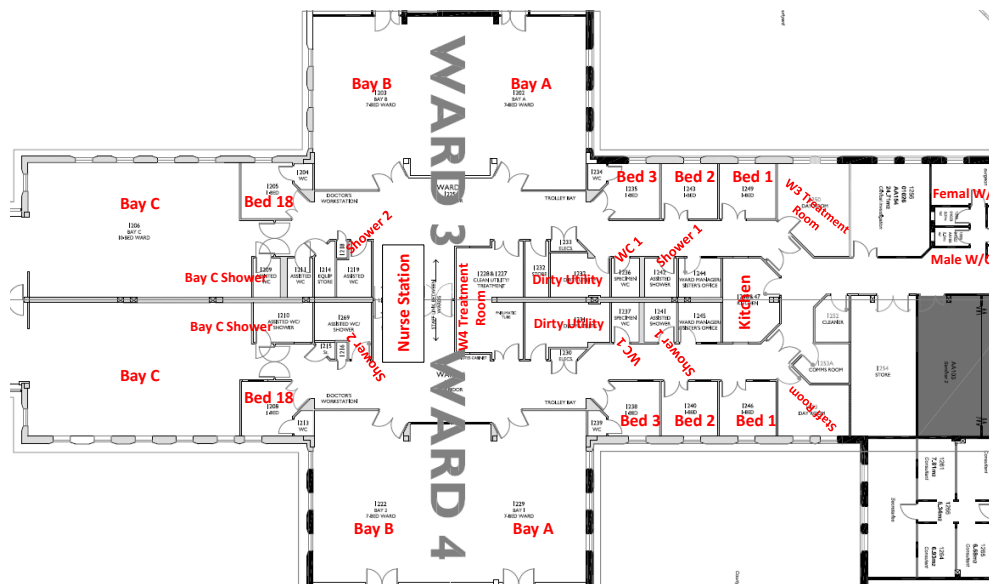
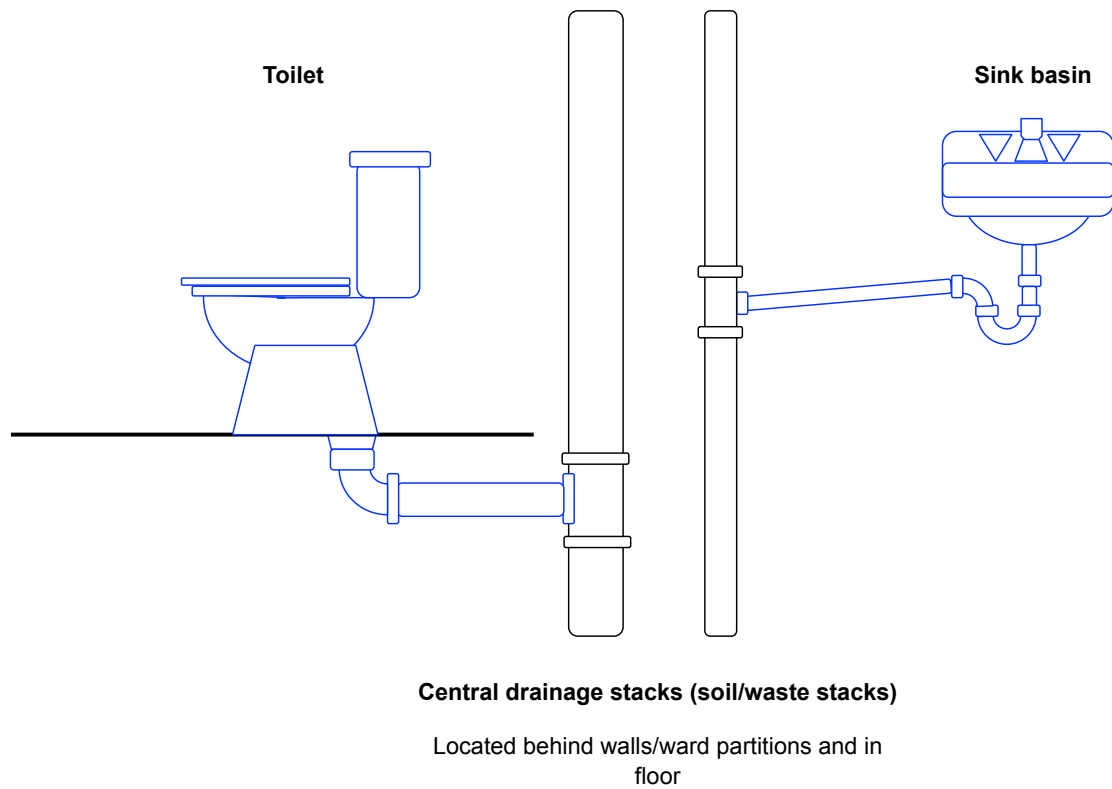


Figure S2. Schematic of sink/wastewater system (not to scale), and extent of plumbing component replacement undertaken as part of control measures (all parts highlighted in blue).



Figures S3-5. Screening history of patients with first CRE detection in the Manchester Heart Centre (MHC) from 01/Apr/2014- 30/Dec/2016. Fig.S2, Fig.S3 and Fig.S4 show individuals with first detection of any CRE, CR-*E. coli* (for which all isolations were from screens), or CR-*K. pneumoniae* respectively. (A) per-patient timelines showing negative screens (dots) prior to first CRE detection. Patients are stratified by year of first CRE detection; (B) histogram showing number of negative screens in the 90 days prior to first CRE detection; (C) interval between last negative screen and first CRE detection.

Figure S3. First CRE detections - any species of CRE

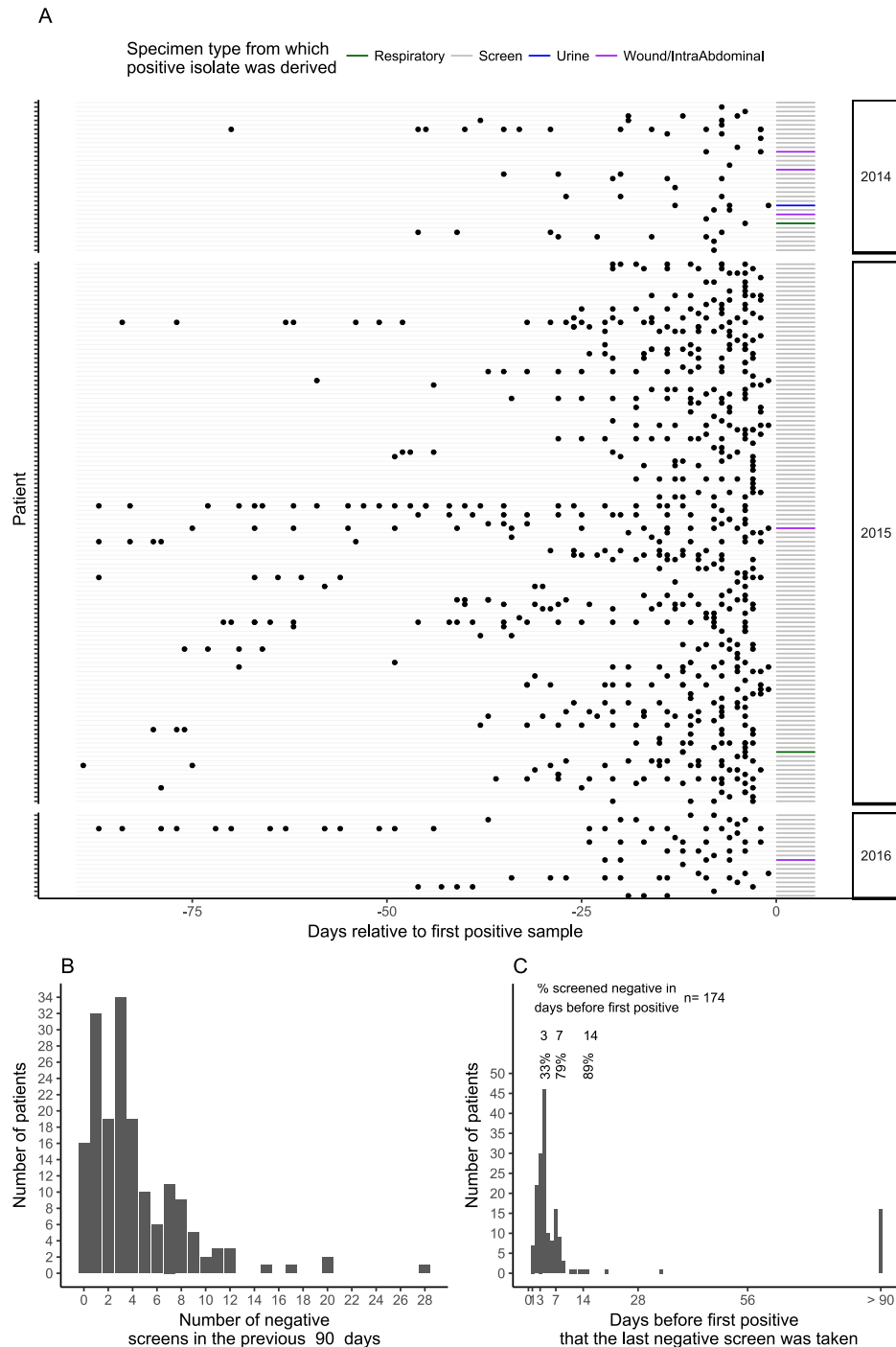


Figure S4. First CRE detections - CR-*E. coli*

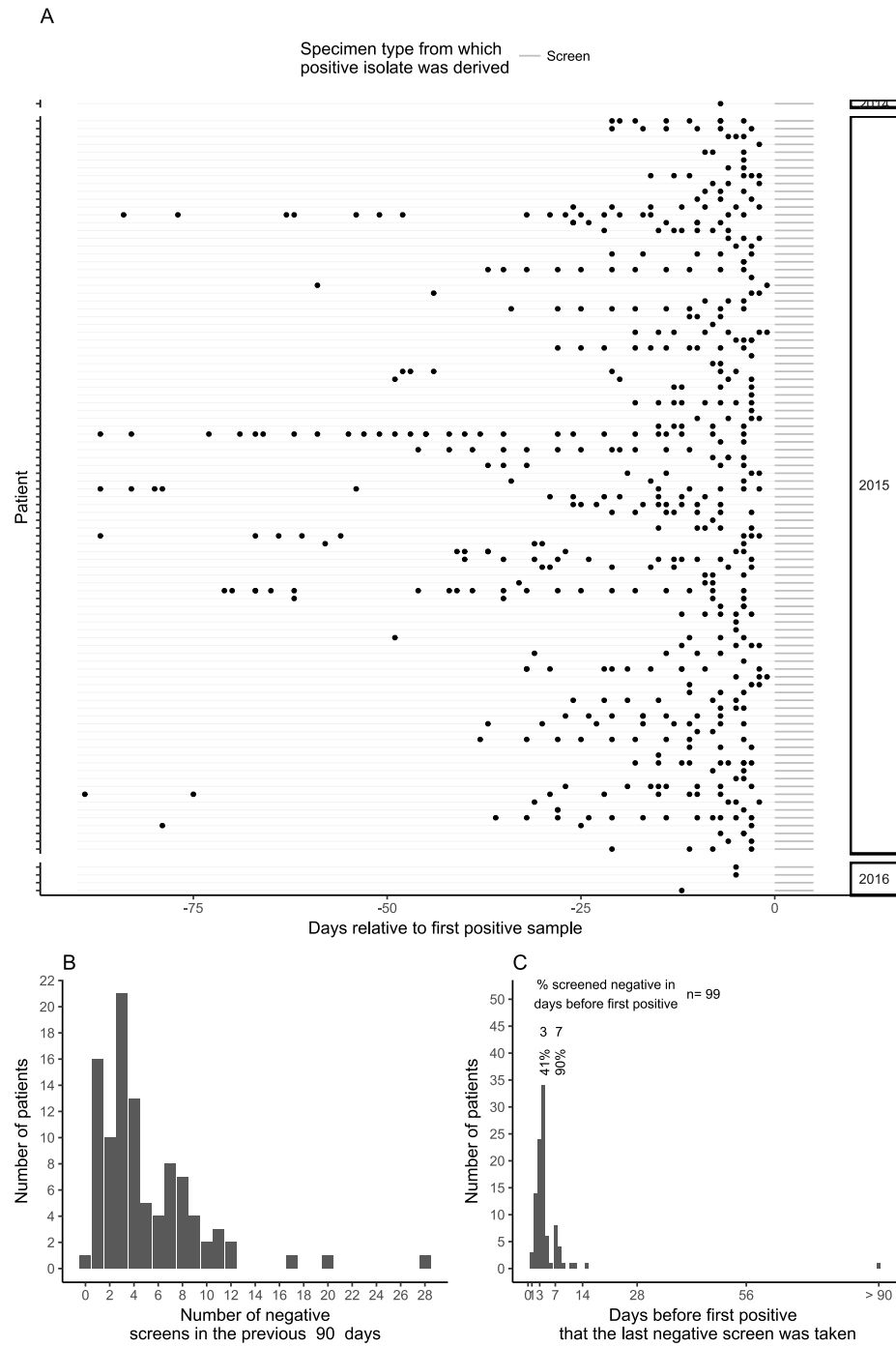


Figure S5. First CRE detections - CR-K. pneumoniae

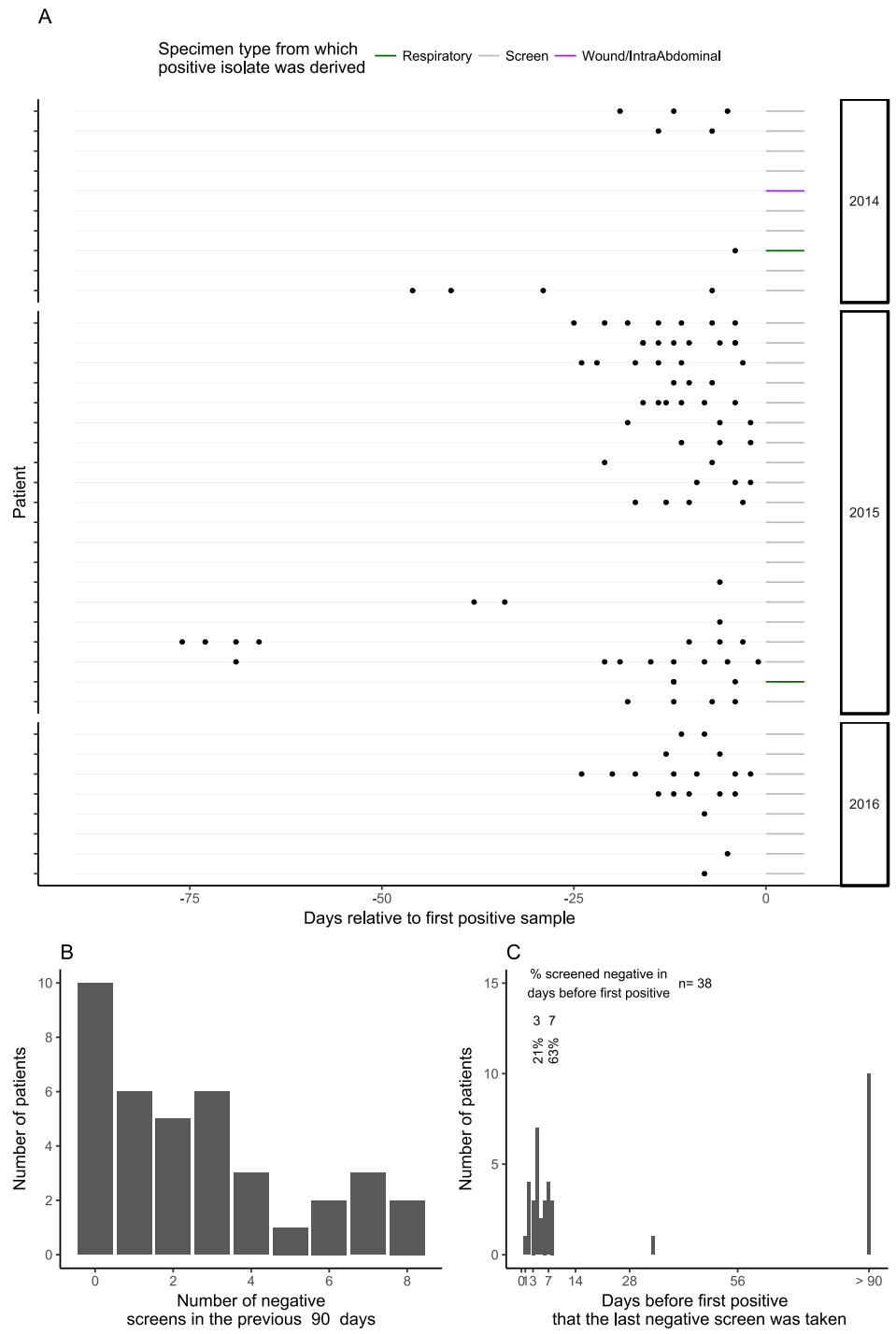


Figure S6. Number of patients screened on \geq day 2 of admission per week by hospital location

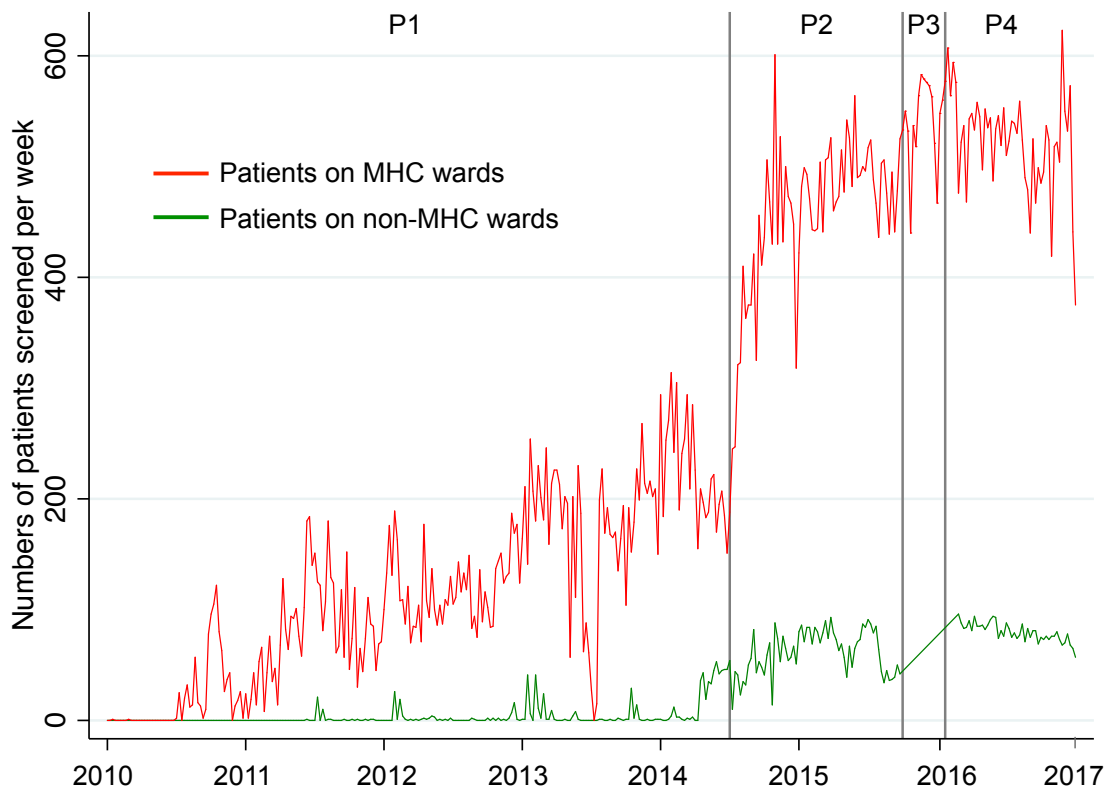


Figure S7. ClonalFrameML output for ST216 outbreak clone KPC-*E. coli* isolates

ClonalFrameML phylogeny of all ST216 KPC-*E. coli* isolates sequenced demonstrating the likely large mega-recombination event (dark blue line) affecting the branch separating ST216 strain-A1 from strain-A2 isolates. “Position” refers to position in genome. Single nucleotide variants were called following mapping to a within-clone derived reference from long-read sequencing for this study.

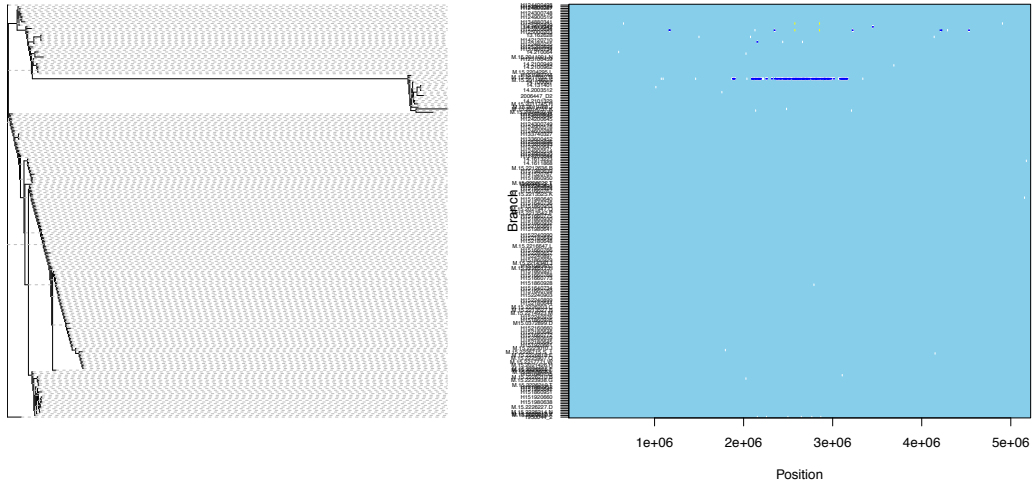
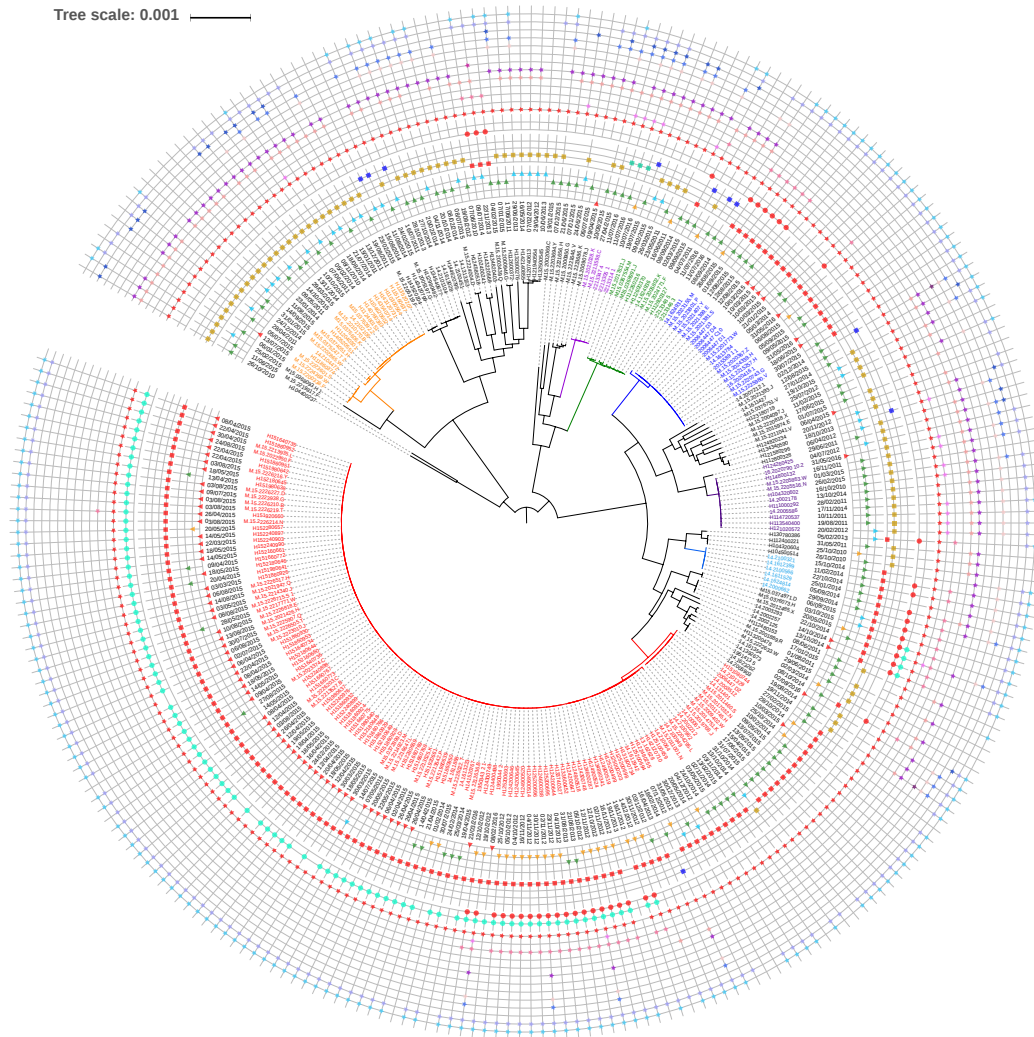


Figure S8. Recombination-corrected KPC-*E. coli* phylogeny with associated antimicrobial resistance gene profiles for 268 sequenced isolates. Coloured stars; from inside out: *bla*_{KPC}, *bla*_{CTX-M-15}, *bla*_{CTX-M-9}, other *bla*_{CTX-M} variants, *bla*_{TEM}, *bla*_{OXA}, *bla*_{SHV}, *bla*_{CMY}, *aac6* variants, *aac3* variants, *aph* variants, *aadA* variants, *qnr* variants). Other annotations are the same as Fig.3 in main text



Supplementary material references

1. Newey WK, West KD. A Simple, Positive Semidefinite, Heteroskedasticity and Autocorrelation Consistent Covariance-Matrix. *Econometrica* 1987; **55**: 703-8.
2. Stoesser N, Sheppard AE, Pankhurst L et al. Evolutionary History of the Global Emergence of the Escherichia coli Epidemic Clone ST131. *MBio* 2016; **7**: e02162.
3. Nguyen LT, Schmidt HA, von Haeseler A et al. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015; **32**: 268-74.
4. Didelot X, Wilson DJ. ClonalFrameML: efficient inference of recombination in whole bacterial genomes. *PLoS Comput Biol* 2015; **11**: e1004041.

5. Bankevich A, Nurk S, Antipov D et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012; **19**: 455-77.
6. Carattoli A, Zankari E, Garcia-Fernandez A et al. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother* 2014; **58**: 3895-903.
7. Stoesser N, Sheppard AE, Peirano G et al. Genomic epidemiology of global *Klebsiella pneumoniae* carbapenemase (KPC)-producing *Escherichia coli*. *Sci Rep* 2017; **7**: 5917.
8. Loman NJ, Quinlan AR. Poretools: a toolkit for analyzing nanopore sequence data. *Bioinformatics* 2014; **30**: 3399-401.
9. Berlin K, Koren S, Chin CS et al. Assembling large genomes with single-molecule sequencing and locality-sensitive hashing. *Nat Biotechnol* 2015; **33**: 623-30.
10. Walker BJ, Abeel T, Shea T et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 2014; **9**: e112963.
11. Kurtz S, Phillippy A, Delcher AL et al. Versatile and open software for comparing large genomes. *Genome Biol* 2004; **5**: R12.
12. Hunt M, Kikuchi T, Sanders M et al. REAPR: a universal tool for genome assembly evaluation. *Genome Biol* 2013; **14**: R47.